Cooperative Design of a Dashboard for Monitoring the P4D Cohort Study on Major Depression

H. Maharlou[1], N. Bössel-Debbert[2], M. Lucht[2], H.B Maier[3], S. Mücke[4], F. Müntefering[5], B. Neuhaus[6], J. Prokein[4]
Ch. Reif-Leonhard[7], J. Voges[5], H. Weber[8], A. Weihs[2], H. Frieling[3] and S. Oeltze-Jafra[1]

[1]Peter L. Reichertz Institute for Medical Informatics, Hannover Medical School, Hannover, Germany
[2]Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Greifswald, Germany
[3]Department of Psychiatry, Social Psychiatry, and Psychotherapy, Hannover Medical School,Hannover, Germany
[4]Hannover Unified Biobank, Hannover Medical School, Hannover, Germany
[5]Institut für Informationsverarbeitung, Leibniz University Hannover, Hannover, Germany
[6]Center for Clinical Trials, Hannover Medical School, Hannover, Germany
[7]Department of Psychiatry, Psychosomatics and Psychotherapy, University Hospital Frankfurt - Goethe University, Frankfurt, Germany
[8]Department of Psychiatry, Psychosomatics and Psychotherapy, University Hospital Würzburg, Würzburg, Germany

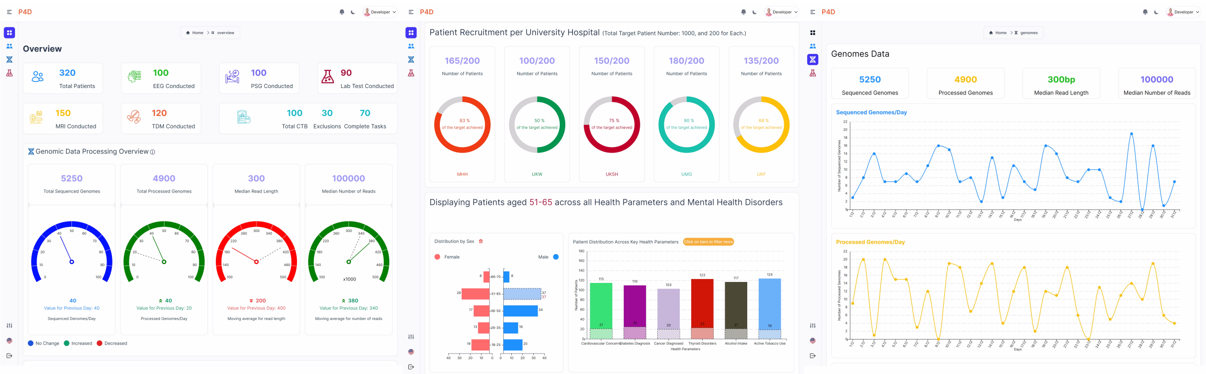

Figure 1: *Current status of the study data acquisition (left), the cohort composition (middle), and the genome sequencing in detail (right).*

Abstract

The P4D (Personalised, Predictive, Precise, and Preventive Medicine for Major Depression) study aims at an improved prediction of treatment outcomes based on a more precise stratification of major depression subtypes. It is collecting very complex data from 1,000 patients across five German university hospitals. We have designed a dashboard to monitor the study and share the collected data among the study partners. We employed a state-of-the-art cooperative dashboard design approach by Setlur et al. [SCST24] in two design cycles: user feedback and dashboard revision. We observed a significant improvement in user satisfaction from the first (Mean=3.57 std=0.95) to the second (Mean=3.87 std=0.80) cycle and an overall positive assessment.

CCS Concepts

• ***Human-centered computing*** → ***Information visualization; Visualization design and evaluation methods;***

1. Introduction

In Europe, Major Depressive Disorder (MDD) affects 8% of individuals annually and 19% over their lifetime. This makes MDD a prevalent mental health issue [KBD*03, SNW22]. It burdens people and healthcare systems, decreasing quality of life and rising disability [dlTVR*21]. Current pharmaceutical interventions for moderate to severe MDD rely on a trial-and-error approach, with 30–50% of patients failing to achieve remission. [MBB22, WWZA23, GRT*08]. Thus, a tailored strategy based on patients' reactions to antidepressants is required. The P4D (Personalised, Predictive, Precise, and Preventive Medicine for Major Depression) study aims at an improved prediction of treatment outcomes based on a more precise stratification of major depression subtypes. It is

DOI: 10.2312/evp.20241081

currently collecting complex data, including omics, magnetic resonance imaging (MRI), polysomnography, laboratory data, and cognitive test results from 1,000 patients at five university hospitals in Germany. We have designed a dashboard to monitor the study's progress and share the collected data among the study partners. The dashboard shall support early detection of acquisition delays, missing data, and biases in cohort composition. Dynamic dashboards are essential for presenting critical data effectively when designed according to professional recommendations for interaction and storytelling. [SCB*19, BS23]. Setlur et al. [SCST24] integrate the Gricean Maxims [Gri75] and Beebe's framework [Bee04]) to promote interactive, cooperative "conversation" between dashboards and users. They provide a set of heuristics for the five conversational states: initiation, grounding, turn-taking, repair & refinement, and closing, termed cooperative dashboard design. We followed Setlur et al.'s approach in designing our dashboard and gathered feedback regarding 39 heuristics from 16 and 12 professional users. We also revised the dashboard based on this feedback in two consecutive design cycles. Our approach and the results are detailed below.

2. Methods

Initial Dashboard Design: Key variables, such as total sequenced genomes, patient recruitment per university hospital, and number of blood samples, were identified from pre-study test data to initiate the dashboard development. A panel comprising seasoned physicians and medical researchers adept in participatory development studies was consulted to curate these key variables. Subsequently, a preliminary dashboard prototype was crafted and evaluated.

Evaluation Instrument and Procedure: A structured survey based on the 39 heuristics from Setlur et al. [SCST24] was created to evaluate the P4D dashboard. These heuristics cover initiation, grounding, turn-taking, repair & refinement, and closing states. Participants rated their agreement with each heuristic on a Likert scale of 1 (strongly disagree) to 5 (strongly agree) with a 'Don't Know' option for challenging or uncertain questions. Furthermore, Open-ended questions were incorporated to obtain comprehensive qualitative input regarding user experiences and possible areas for enhancement. The evaluation was conducted in two rounds, before and after dashboard revision, involved 16 and 12 prospective users testing the dashboard autonomously.

Survey Data Analysis: *Quantitative Analysis*: Descriptive statistics (mean scores and standard deviations) were computed for each round and heuristic. Heuristics were categorised according to participant agreement into low/moderate/high agreement, i.e. <50%/50-75%/>75% of participants assigned a score of 4 or 5 to the heuristic. *Qualitative Analysis*: Thematic analysis was used for open-ended comments to find reoccurring themes regarding the use of heuristics, dashboard strengths, and particular design difficulties.

Dashboard Design Revision: We revised the dashboard by focusing on heuristics (H) with low and moderate agreement and addressing open-ended comments. For instance, H3 and H11 (see appendix for details) initially received moderate ratings due to unclear instructions, undefined interaction points, and a lack of visual cues such as icons. Subsequently, we implemented enhancements by introducing notifications that appear when users scroll down and reach interactive charts, thereby indicating the possibility of interaction. The dashboard also employs intuitive conventions, ensuring that icons and symbols are easily comprehensible for users. For instance, icon placement and appearance semantics effectively communicate patterns in the data, such as adjusting arrow icons to signify ascending or descending sorting in tables. These modifications significantly improved comprehensibility and interaction, significantly enhancing both heuristics' ratings from 3.33 to 4.45 and 3.73 to 4.17, respectively. Parts of the revised dashboard are shown in Figure 1. The dashboard can be tested with dummy data at `https://p4d.vercel.app/`.

3. Results

Out of 39 heuristics, 11 (28%) were improved from moderate and low agreement to high agreement after addressing the comments of the first evaluation round (Figure 2). The number of moderate agreements decreased from 22 (56%) to 14 (35%) in the second round, the same for low score heuristics as they became 5 (12.8%) in the second round compared to 8 in the first round (20%). The level of agreement per conversational state after dashboard revision is shown in Figure 3. On average, initiation and grounding show the highest level of agreement, whereas turn-taking, repair & refinement and close offer most room for improvement.

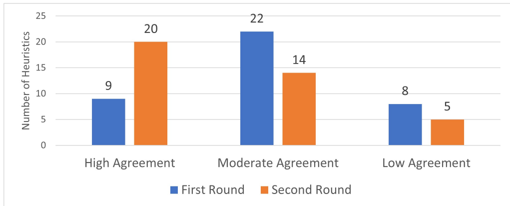

Figure 2: *High, moderate, and low agreement heuristics for the first and second evaluation rounds.*

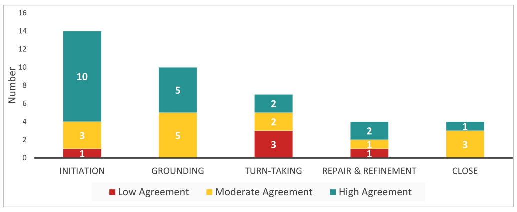

Figure 3: *The number of heuristics in low, moderate, and high agreement for the five conversational states.*

4. Conclusions

The cooperative design and evaluation of the dashboard yielded an average agreement of 3.87 with the 39 heuristics proposed by Setlur et al. [SCST24]. In future work, we will focus on improvements regarding the 5 remaining heuristics with poor agreement. We will, for instance, enable dashboard customization to improve agreement with H35

References

[Bee04] BEEBE S. A.: *Interpersonal Communication: Relating to Others*. Pearson Education Canada, 2004. 2

[BS23] BURCH M., SCHMID M.: *Dashboard design*. River Publishers, 2023. URL: https://www.routledge.com/Dashboard-Design/Burch-Schmid/p/book/9788770040044, doi:https://doi.org/10.1201/9781032657301. 2

[dlTVR*21] DE LA TORRE J. A., VILAGUT G., RONALDSON A., SERRANO-BLANCO A., MARTÍN V., PETERS M., VALDERAS J. M., DREGAN A., ALONSO J.: Prevalence and variability of current depressive disorder in 27 european countries: a population-based study. *The Lancet Public Health 6* (10 2021), e729–e738. doi:10.1016/S2468-2667(21)00047-5. 1

[Gri75] GRICE H. P.: *Logic and Conversation*. Broadview Press, 1975, p. 47. 2

[GRT*08] GAYNES B. N., RUSH A. J., TRIVEDI M. H., WISNIEWSKI S. R., SPENCER D., FAVA M.: The star*d study: treating depression in the real world. *Cleveland Clinic journal of medicine 75* (2008), 57–66. URL: https://pubmed.ncbi.nlm.nih.gov/18236731/, doi:10.3949/CCJM.75.1.57. 1

[KBD*03] KESSLER R. C., BERGLUND P., DEMLER O., JIN R., KORETZ D., MERIKANGAS K. R., RUSH A. J., WALTERS E. E., WANG P. S.: The epidemiology of major depressive disorderresults from the national comorbidity survey replication (ncs-r). *JAMA 289* (6 2003), 3095–3105. URL: https://doi.org/10.1001/jama.289.23.3095, doi:10.1001/jama.289.23.3095. 1

[MBB22] MINELLI A., BARLATI S., BAUNE B. T.: Evaluating study designs and treatment outcomes of antidepressant pharmacogenetic clinical trials - challenges and future perspectives. a critical review. *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology 59* (6 2022), 68–81. URL: https://pubmed.ncbi.nlm.nih.gov/35561539/, doi:10.1016/J.EURONEURO.2022.04.007. 1

[SCB*19] SARIKAYA A., CORRELL M., BARTRAM L., TORY M., FISHER D.: What do we talk about when we talk about dashboards? *IEEE Transactions on Visualization and Computer Graphics 25* (1 2019), 682–692. doi:10.1109/TVCG.2018.2864903. 2

[SCST24] SETLUR V., CORRELL M., SATYANARAYAN A., TORY M.: Heuristics for supporting cooperative dashboard design. *IEEE Transactions on Visualization and Computer Graphics 30* (1 2024), 370–380. URL: https://doi., doi:10.1109/TVCG.2023.3327158. 1, 2

[SNW22] SHOREY S., NG E. D., WONG C. H.: Global prevalence of depression and elevated depressive symptoms among adolescents: A systematic review and meta-analysis. *The British journal of clinical psychology 61* (6 2022), 287–305. URL: https://pubmed.ncbi.nlm.nih.gov/34569066/, doi:10.1111/BJC.12333. 1

[WWZA23] WANG X., WANG C., ZHANG Y., AN Z.: Effect of pharmacogenomics testing guiding on clinical outcomes in major depressive disorder: a systematic review and meta-analysis of rct. *BMC psychiatry 23* (12 2023). URL: https://pubmed.ncbi.nlm.nih.gov/37173736/, doi:10.1186/S12888-023-04756-2. 1

Proceedings published by Eurographics - The European Association for Computer Graphics.